



地理的犯罪予測手法の改良

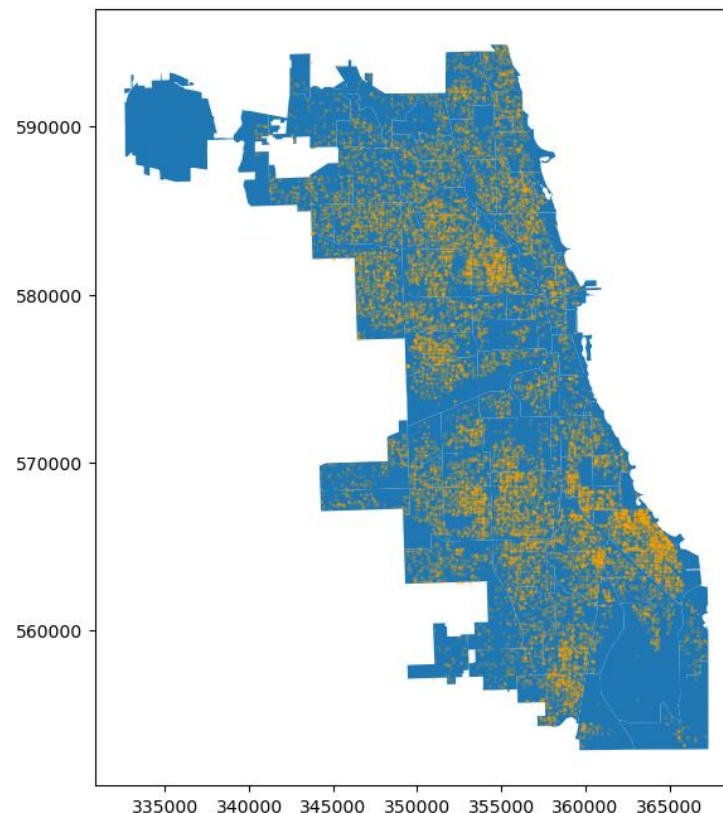
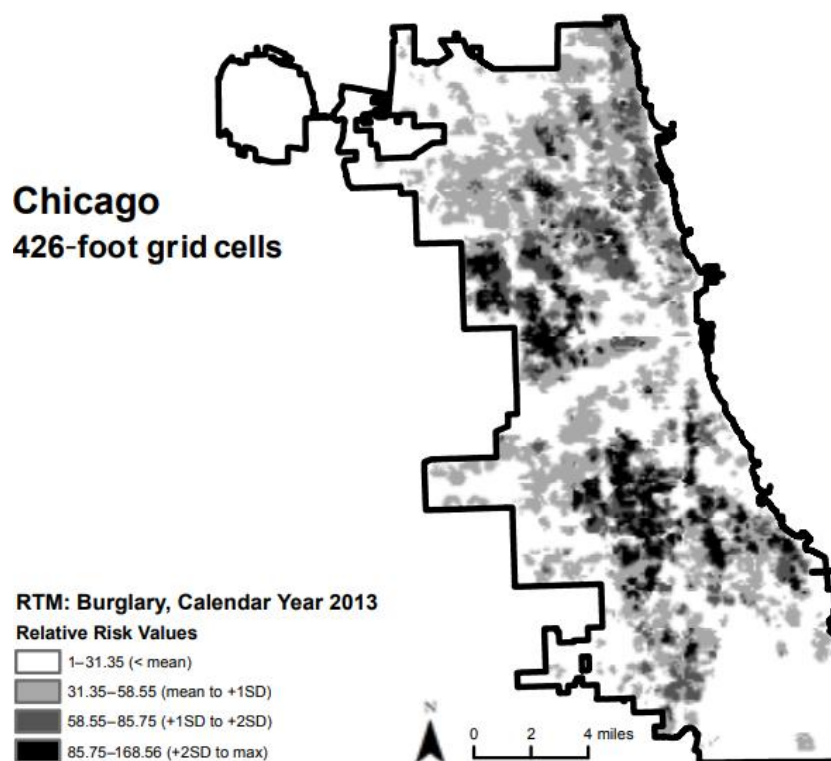
木脇研 B4 橋本響



■ 研究背景：RiskTerrainModeling(RTM) [J. M. Caplan+ 15]

・地理情報を活用して犯罪を予測する手法

- 地理的要因から対象領域の犯罪リスクを推定するモデル
- 視覚的に高リスクエリアと低リスクエリアを表現
- 犯罪に関連する地理的要因の重要度を分析



出典:[J. M. Caplan+ 15]

■ 研究背景：RTMのアルゴリズム

- 対象領域をグリッド状に分割
- 犯罪に関連する地理的要因からレイヤーを生成
- 地理的な要因ごとの犯罪に関連する重みを算出
- 各レイヤーに重みをかけて重ね合わせてリスクを求める

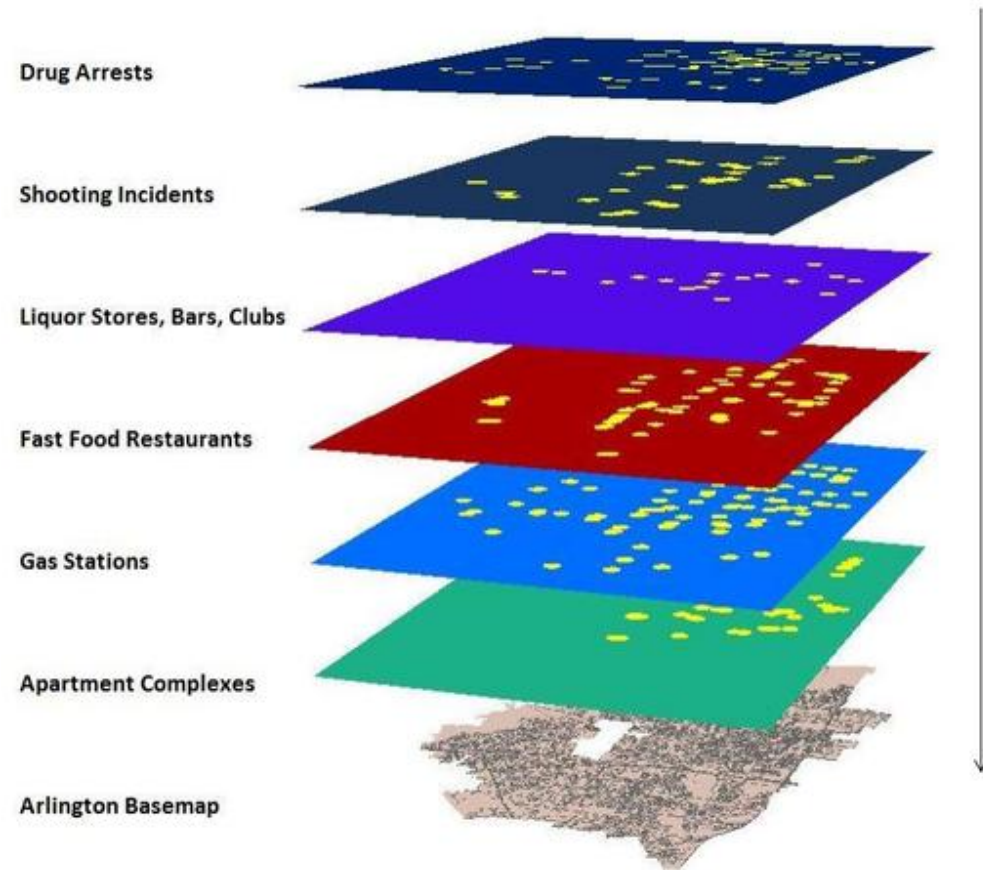
地理的要因を予測変数
犯罪件数を目的変数

$$\text{強盗} = \beta_0 + \beta_1 \text{放置車両} + \beta_2 \text{廃墟} + \dots$$



変数選択

β を決定



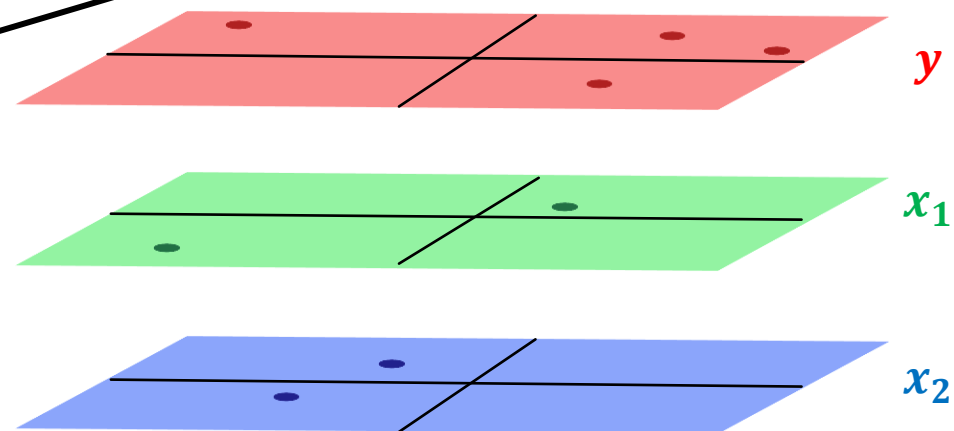
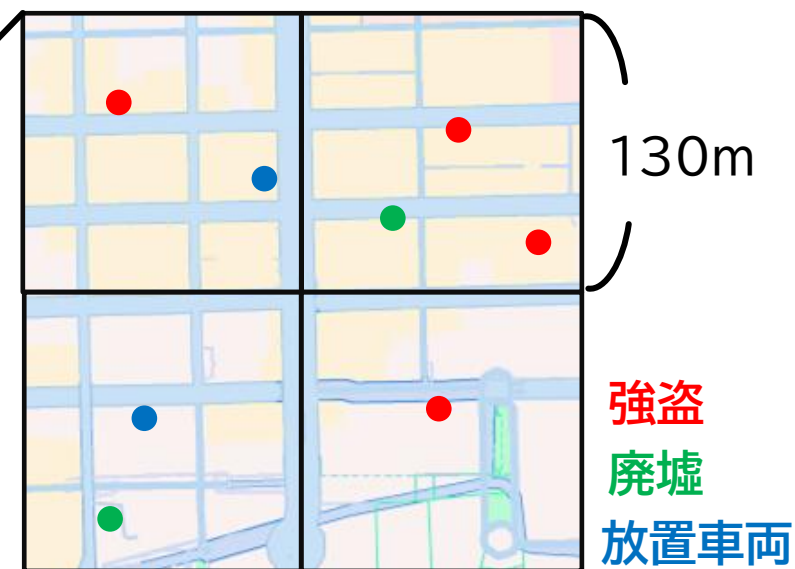
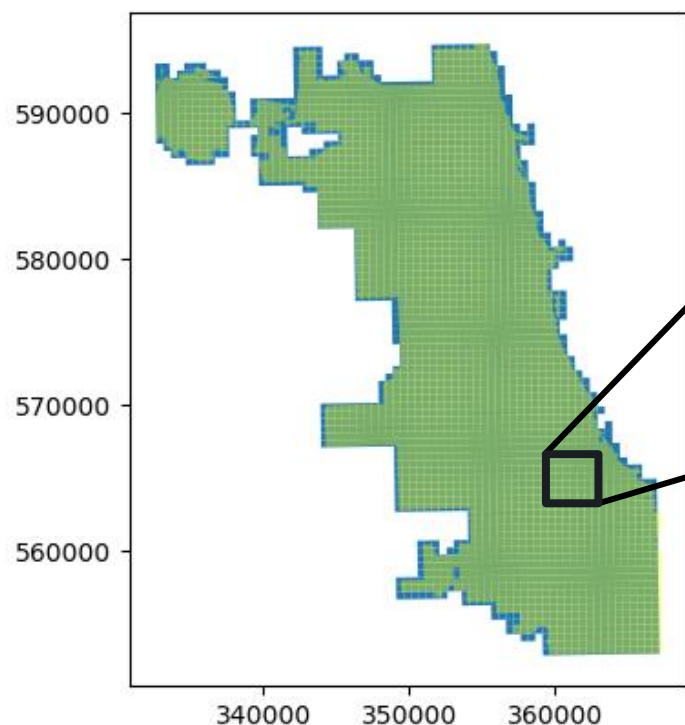
出典:<http://stephenmaps.weebly.com/analysis.html>

■ 研究目的

- RTMを**予測精度**の観点から改良
 - RTMを実装
 - 探索的データ分析によりモデル改善
- RTMを**解釈性**の観点から改良
 - 地理的要因と犯罪の因果関係を推計する
 - 統計的因果推論を活用

■ 研究背景：レイヤーの構成

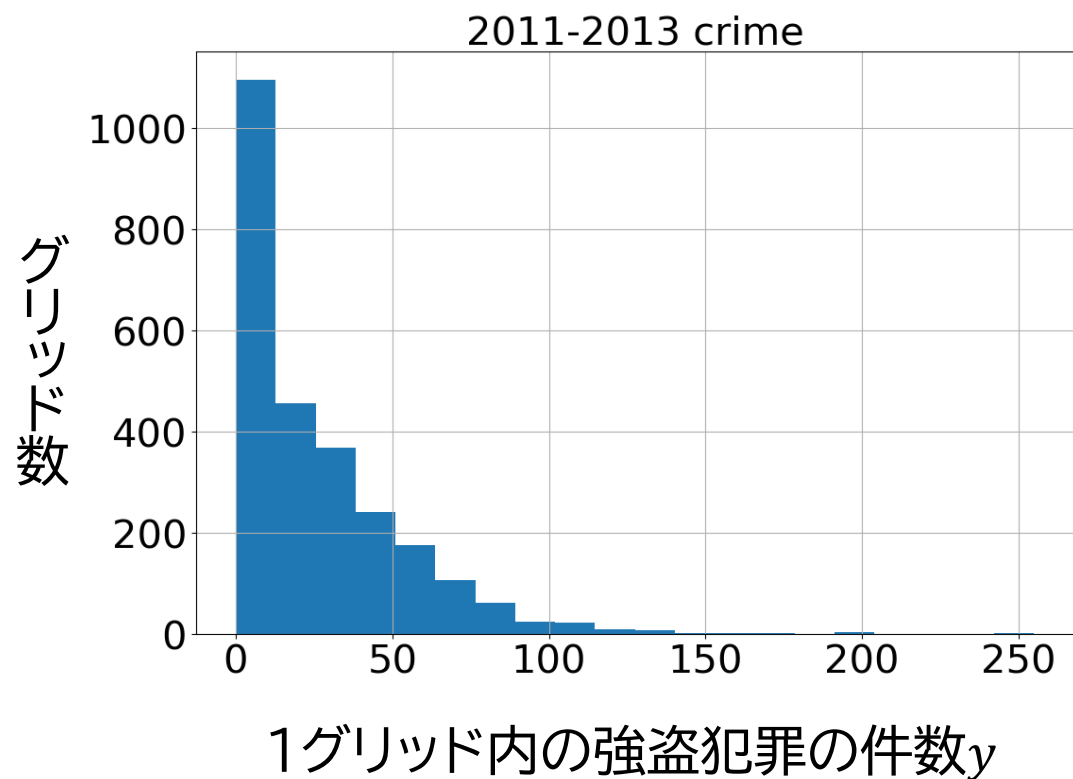
- ・オープンデータが豊富なシカゴを対象
- ・Chicago Data Portal^[CDPH,24] を利用
- ・シカゴのグリッドマップを作成
- ・グリッド幅は130m×130m



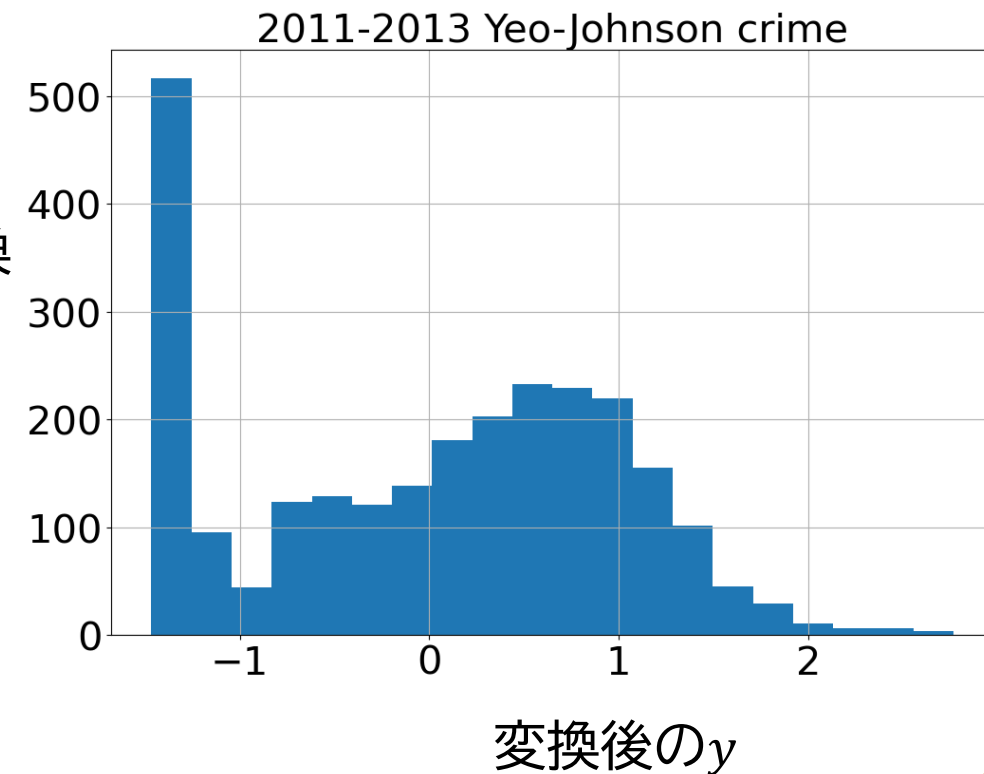
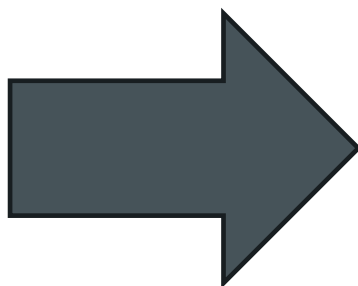
■ 進捗：RTMの実装

①データ前処理

- Yeo-Johnson Power Transform^[Weisberg 01]
- 犯罪の分布が**正規分布**に近くなるよう変換



yeo-johnson変換

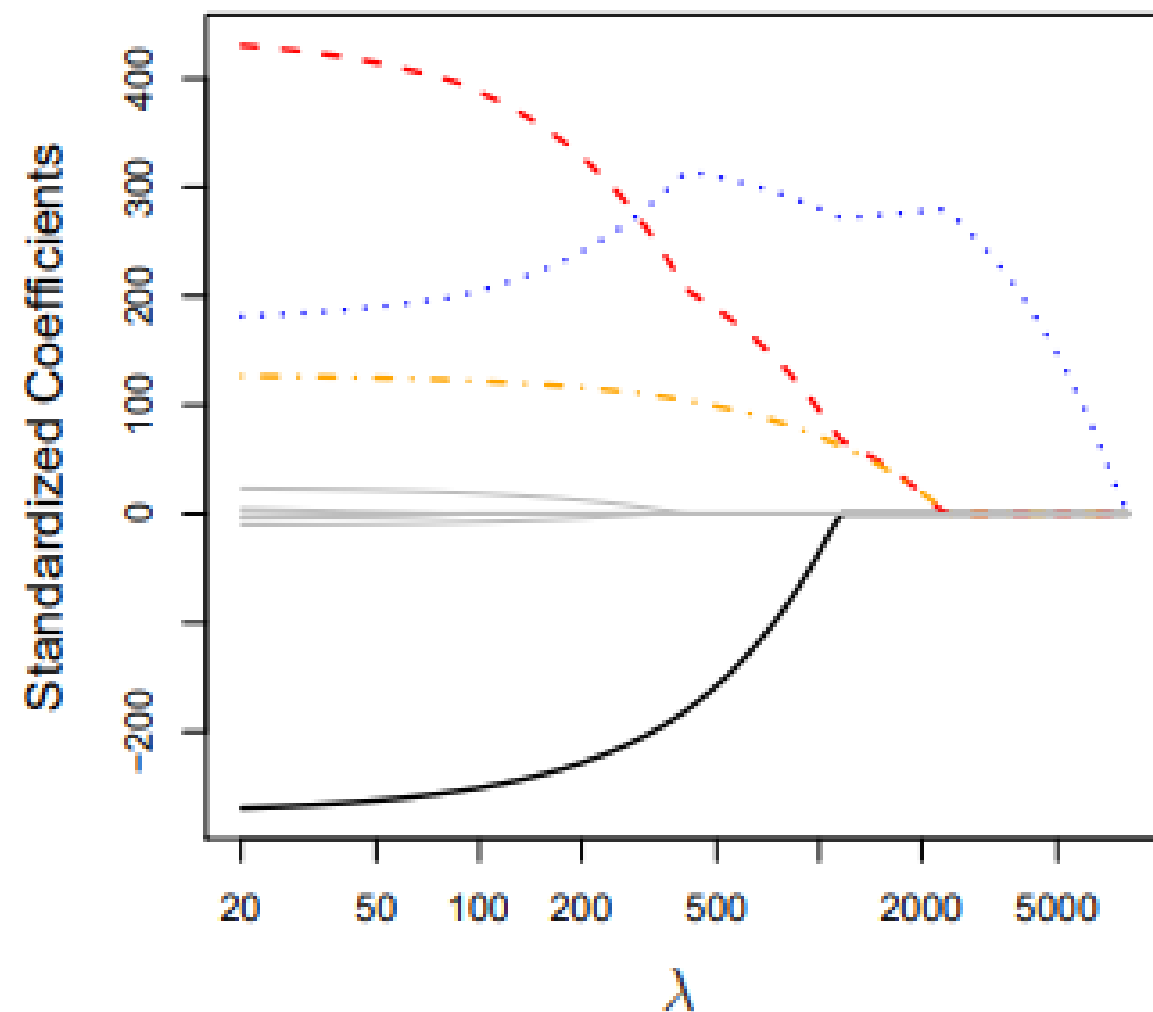


■ 進捗：RTMの実装

② Lasso回帰で変数選択^[James+ 23]

- ペナルティパラメータ λ は自分で設定
- 20fold交差検証^[James+ 23]で λ を決定
- オリジナルのRTMの変数選択を再現

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j|$$



出典:[Gareth James+ 23]

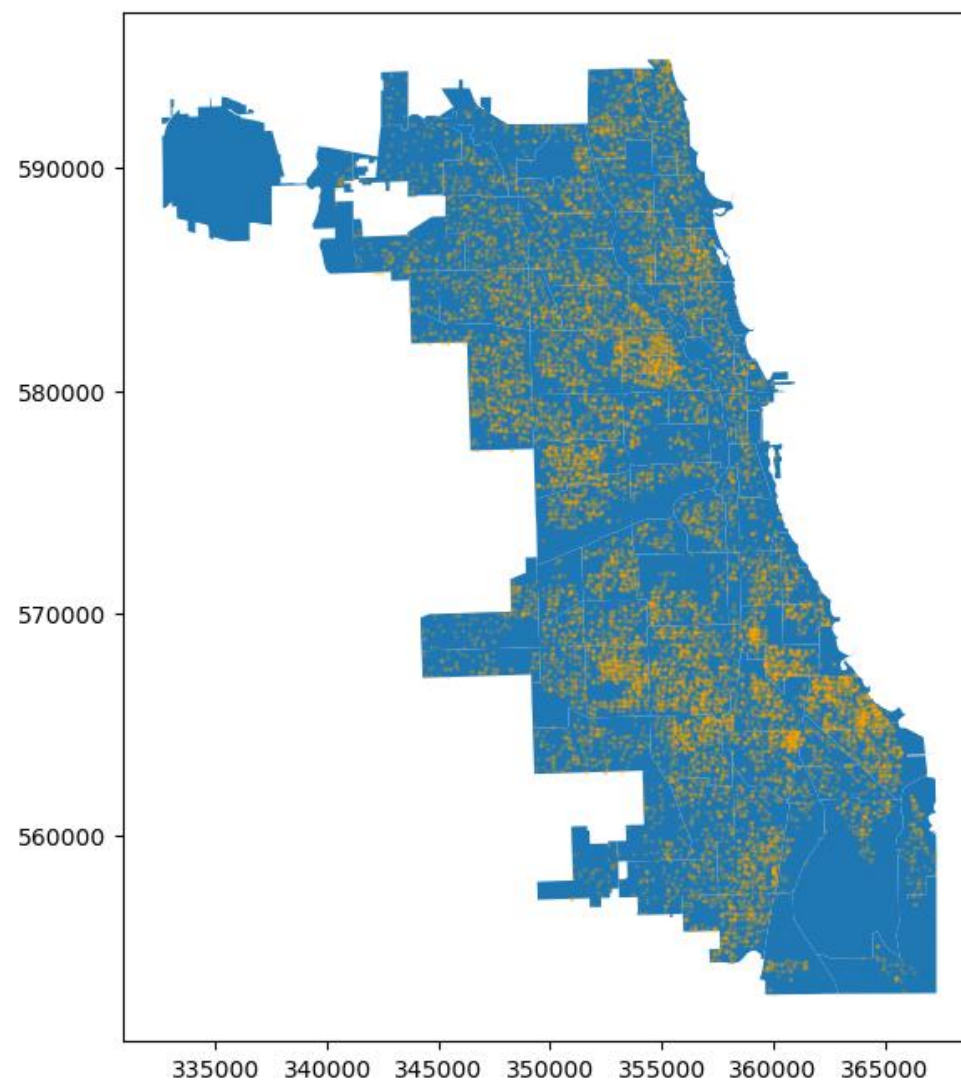
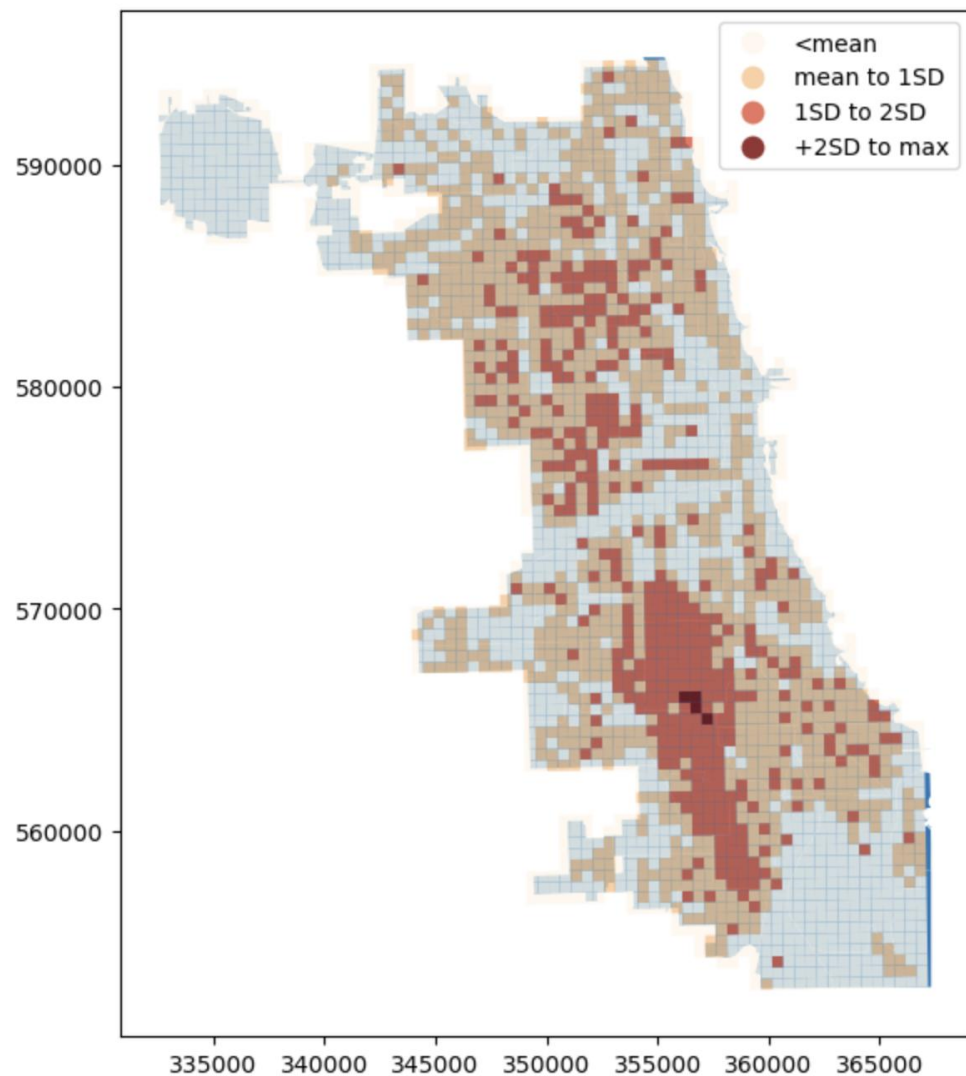
■ 進捗：実験条件

- ・学習データ：2011年～2013年
- ・テストデータ：2014年

	2011	2012	2013	2014
強盗	26,615	22,832	17,885	14,560
放置車両	19,933	17,411	16,091	20,385
廃墟	15,380	11,976	8,366	5,450
木柱街灯の消灯	46,844	20,025	15,221	21,762
金属柱街灯の消灯	34,059	31,033	22,748	65,155
学校	674	681	672	680
差し押さえ物件	16,680	16,120	11,131	7,511
落書き	136,948	109,931	137,079	124,742
不衛生な場所	17,920	19,122	18,092	19,042

■ 進捗：学習結果

・2014年の犯罪リスク値を図示



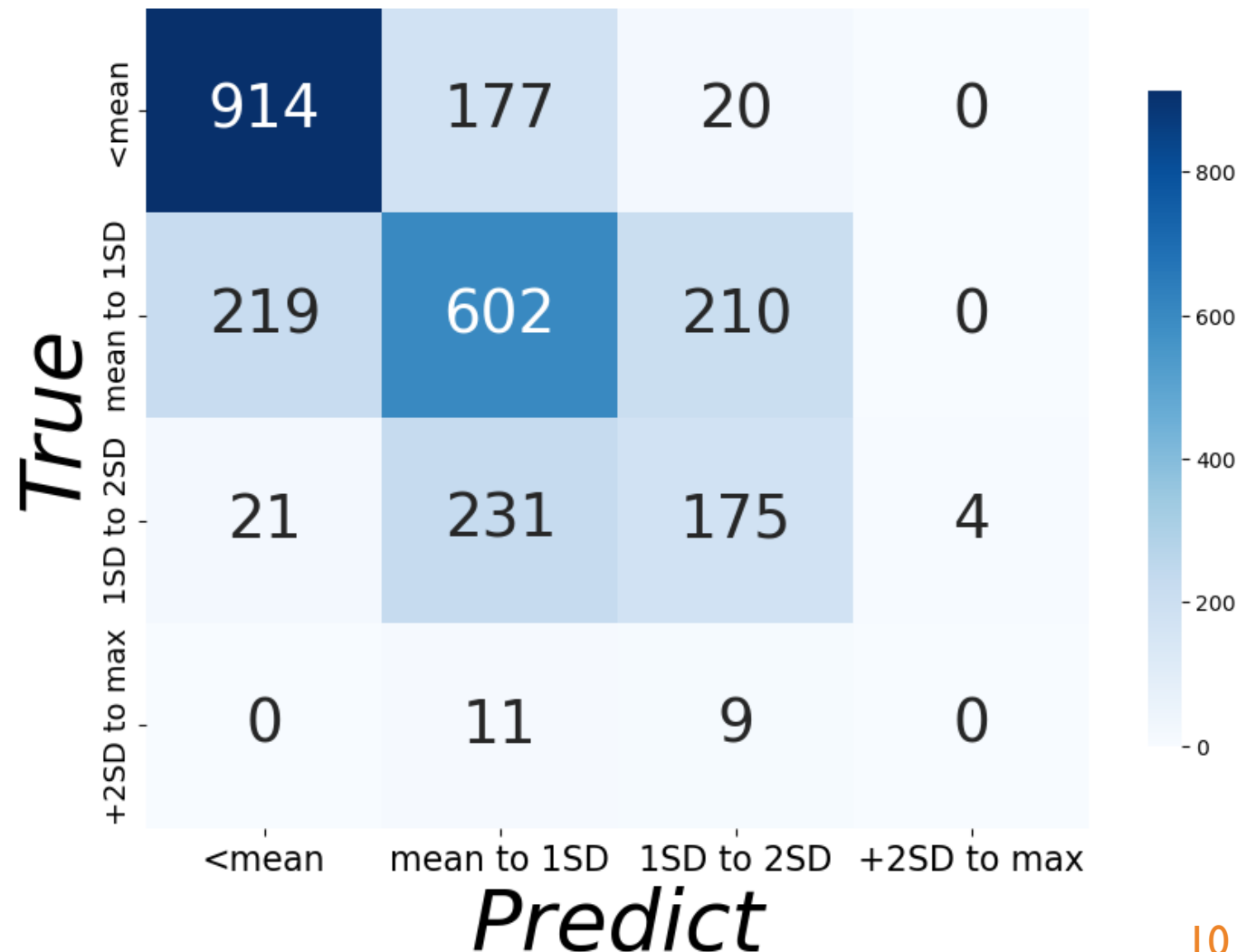
■ 進捗：学習結果

- 2014年の犯罪リスク値と犯罪件数を比較

- 全てのグリッドを4クラスに分割

- ~平均
- 平均~1標準偏差
- 1標準偏差~2標準偏差
- 2標準偏差~

- 対角線のグリッドは正しく予測
- 約65%の予測精度



■ 展望

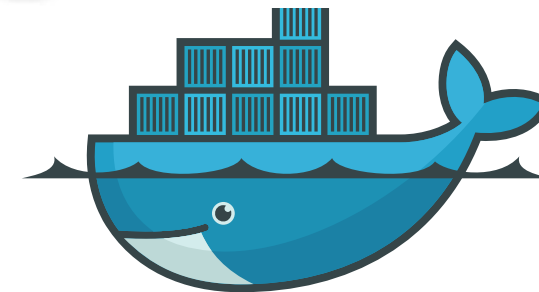
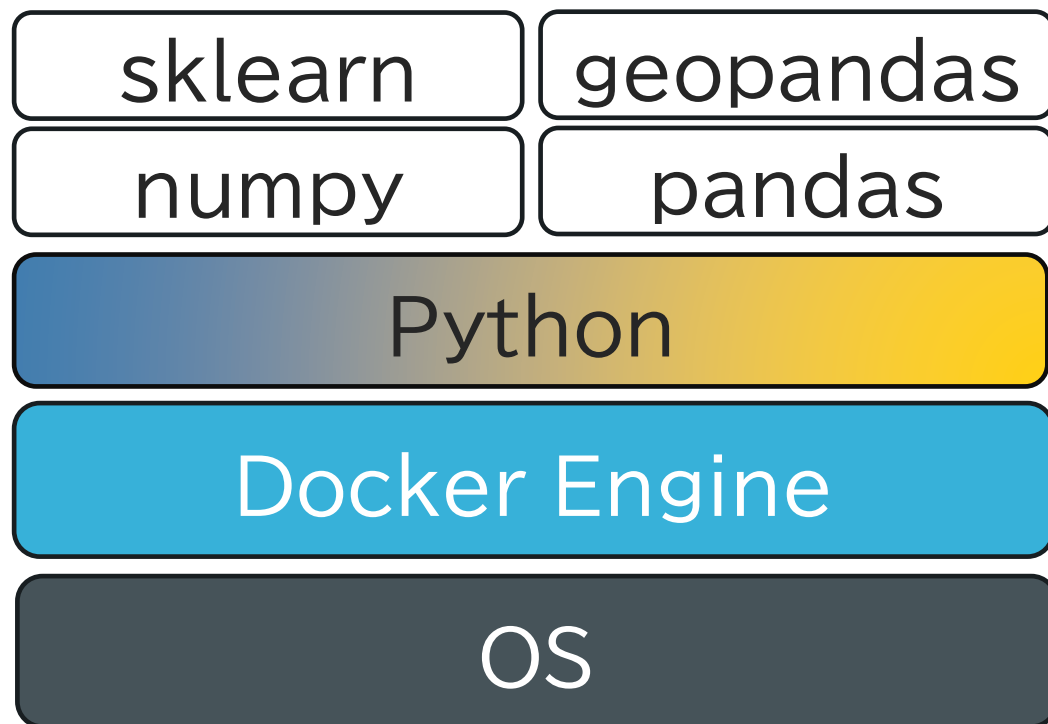
- 住所に基づく位置情報も利用できるようにする
- 犯罪間の**時間的・地理的**な相関関係を調査
- 統計的**因果推論**を活用した犯罪要因の因果関係を推計
- モデルの取りこぼしを細かく調査

■ 文献リスト

- [Caplan+ 15] Caplan, J., Kennedy, L., and Barnum, J. Risk Terrain Modeling for Spatial Risk Assessment, 2015
- [CDPH, 24] Chicago Department of Public Health. Chicago Data Portal, (2024) (<https://data.cityofchicago.org/>)
- [James+ 23] Gareth James, Daniela Witten, Trevor Hastie, Rob Tibshirani, Jonathan Taylor. *An Introduction to Statistical Learning with Applications in Python*, 2023
- [Weisberg 01] Sanford Weisberg. Yeo-Johnson Power Transformations, 2001

■ 実行環境

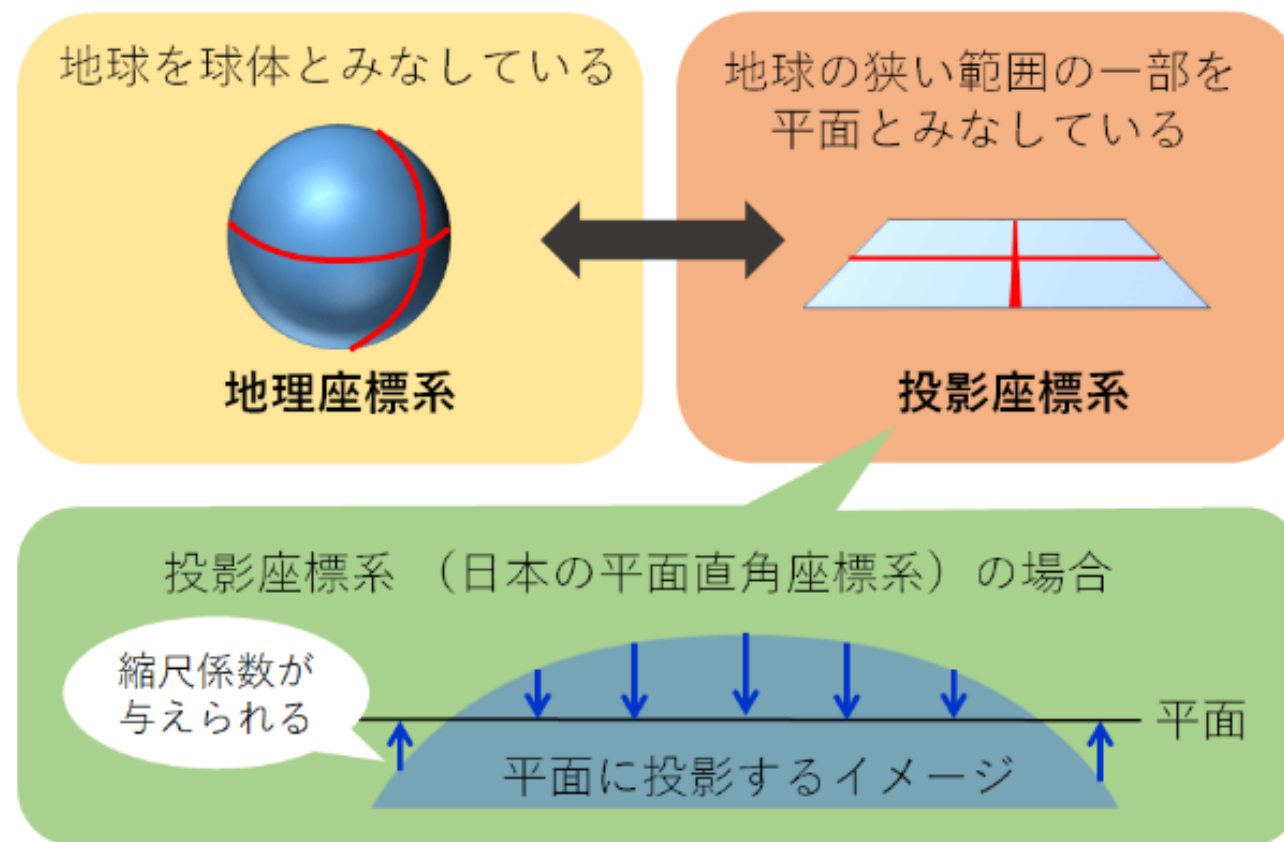
- Dockerを利用した仮想環境



docker

■ 研究手法

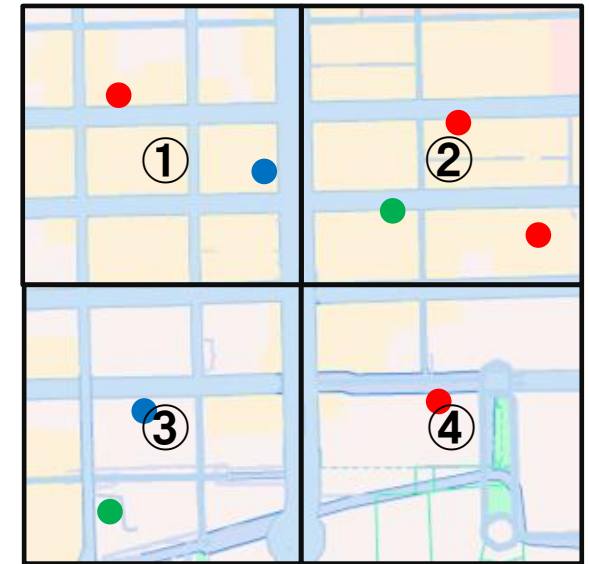
- ・座標系の変換
 - 緯度経度から距離を扱える座標系に



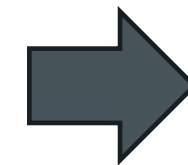
■ 研究手法

- ・犯罪に関連する地理的な要因の位置情報を**予測変数**
- ・各グリッドの犯罪件数を**目的変数**

	強盗	放置車両	廃墟	...
グリッド①	0	1	0	
グリッド②	2	0	1	
グリッド③	0	1	1	
グリッド④	1	0	0	
⋮				



$$\text{強盗} = \beta_0 + \beta_1 \text{放置車両} + \beta_2 \text{廃墟} + \dots$$



β を決定